Precipitation Pattern in the Western Himalayas revealed by Four Datasets

Hong Li, Jan Erik Haugen, Chong-Yu Xu

HESS manuscript: **hess-2017-296**

Recommendation: **Reconsider after major corrections**

This paper provides a limited examination of four precipitation datasets over the Western Himlayan region of India, and is primarily performed by comparing the spatial pattern and trends of the datasets to identify which datasets are suitable for hydrological modelling. The results are short, but there is a conclusion that WRF and APHRODITE are useful tools. While the paper address an important issue faced by hydro-meteorologists, there are still major problems in the strength and sufficiency of the methodology and of the analyses. While there are improvements in this revised version, there are major issues in the justification of the methodology and the analysis of the results, as well as a lack of detail regarding the WRF modelling hinders the replicability of the results. In my opinion, the current manuscript does not offer sufficiently thorough support to amount to a substantial advancement of understanding precipitation patterns in this region. Therefore, I recommend that the manuscript should be reconsidered for publication after major revisions.

Major comments:

1. The major limitation of this manuscript is that there is still a lack of benchmarking to determine which is the "best" dataset for use in hydrological modelling. While the discharge and runoff measurements are a step in this direction, I have to question why the authors do not simply use *in situ* observations as the benchmark, and compare the four data sets to this data. Daily and monthly observed precipitation is freely available from the National Climatic Data Center (https://gis.ncdc.noaa.gov/maps/ncei/cdo/daily) for this region. It would be prudent for this data to be included in this manuscript, given that there is currently no direct benchmark for precipitation used. Although the authors identify the lack of precipitation measurements and their quality is limited in the Himalayan region, including this data and the MODIS data as the benchmark will dramatically improve the veracity of the analysis and conclusions. This will also help determine what added value the "best" dataset has over the others.

2. As well as the spatial pattern and trends in precipitation, whether or not extreme precipitation events are captured by the individual data sets would also be of value to hydro-meteorologists. The "best" data set cannot be identified just from the mean spatio-temporal characteristics, when extreme events are important for accurate hydrological modelling.

3. A missing component of the work as a whole is that there is no proven identification of the real source of the differences. The differences are just explained, without any digging into the simulated processes and forcing data to really find the causes. The key to a valuable comparison study is to at least isolate the source of the differences so that others may understand and build on the work. While the authors identify that differences between the data sets may be attributed to differences in grid size resolution, why not regrid the four data sets onto the same grid and then compare and analyse? On line 26, line 5, the authors state that horizontal resolution is the reason for APHRODITEs better performance. Is this true, or is it an artefact of the individual observations? The authors later attribute ERA-Interim's poor performance to individual observations, rather than horizontal resolution. So which is it, horizontal resolution or an artefact of the number/distribution/quality of observations? This is contradictory and confusing and does not isolate the true reason for differences between data sets.

4. A big question is whether the differences between the data sets are statistically significant. The absence of this is a real deficiency and should be addressed. The reader just doesn't know whether the small differences (e.g. in the spatial differences) have any statistical meaning and which of the different data sets are truly better. It may be that it cannot be concluded that any of these are statistically significantly different. But, even to know that would be of value. Without significance testing results, the reader can't conclude either way.

5. Some very crucial information is required in the WRF section to clarify how the simulations were conducted. It is the only dataset to have been produced by the authors, and the scant detail currently given means that there is no way of replicating their results. Details that need to be clarified in the text are as follows:
    a. What justification is there for using such a precise resolution of 16306 km, rather than 16 km? Maussion *et al.* and Li *et al.* are cited for their model setups, despite both studies using a nested model approach (with the innermost domain less than 3 km). The authors have assumed that model performance (of the other two studies) will scale appropriately to 16306 km. What sensitivity simulations were conducted to test for this? Since this region is surrounded by complex topographic features, and the aim of this paper is to emphasise the benefit of WRF over the other coarser gridded data sets, it would seem appropriate to use a nested model approach in this study going to a resolution of > 5km.
    b. The authors state that it is not easy to determine the optimised selection of parameterization schemes – this is correct, and setting the model up for a region often involves intensive testing of the physics options and sensitivity simulations. So please justify why some options from Maussion *et al.* and some options from Li *et al.* have been selected without proper testing? This is contradictory, and there is no adequate justification for doing this (just use one setup or the other).
    c. What was the height of the lowest vertical level (usually ~25m)?
    d. Did the authors consider using other reanalysis products (e.g. MEERA, GFS)?

e. Is there any erroneous data at the boundary edges that run through the Himalayas?
f. What topography data set was used in WRF? Did the authors compare the modelled topography to the real observed topography? Even good elevation agreement, such as being within 10m, can affect the results at the precision reported.
g. At no point in this section have the authors explained how the data has been extracted from WRF. Was the data extracted from the nearest grid point or an interpolation?
h. Were these runs continuous, restarted, or reinitialised, and what was done to check for model spin up and drift? It is really important to make this clear.
i. Where did the SST data come from, and how frequently was this updated?

6. There is still a lack of verification indicators, and including the data from available rain gauges could easily be done. As highlighted above, WMO rain gauge data is freely available. Daily precipitation may also be available from the Bhakra Beas Management Board (see e.g. Norris *et al.* 2017)

Norris, J., Carvalho, L. M., Jones, C., Cannon, F., Bookhagen, B., Palazzi, E., & Tahir, A. A. (2017). The spatiotemporal variability of precipitation over the Himalaya: evaluation of one-year WRF model simulation. *Climate Dynamics*, *49*(5-6), 2179-2204.

Other comments:

7. I recommend that the authors carefully proofread the manuscript again. There are numerous grammatical and typographical, none of which I have corrected. Occasionally, the clarity of the arguments is lost by poor sentence structure. This makes the manuscript hard to follow at times.
8. Revise lines 5 – 9 on page 2, the paragraph is hard to follow
9. Although the weaknesses are discussed, I would like to see the advantages of each of the datasets also included in Table 1, as well as in the corresponding paragraph, for a two-sided comparison.
10. Line 5, page 3 – "This interaction brings plenty of precipitation" – This is a qualitative statement which is not useful here. Use the WMO data to state exactly how much; be precise.
11. Lines 16 – 20, page 3 – What is the relevance of this paragraph in this section? What is happening to the glacier? Consider removing, or expand to make relevant.
12. Have you considered also including TRMM (3B42V7) as well as/instead of APHRODITE in the analysis? TRMM is the most reliable decadal dataset of gridded precipitation estimates in the Himalaya (Norris *et al.*), and APHRODITE is mostly based on low elevation sites.
13. Explicitly state where the discharge measurements were sourced. This information appears to be missing.
14. Lines 10 – 12, page 6 – This detail is more appropriate in the method section

15. Lines 16 – 21, page 6 – this paragraph does not make sense at all. Be concise and analytical.
16. Lines 3 – 7, page 7 – This section should also be in the method section. If you are not using *in situ* observations as the benchmark for "best" then use this as your justification in the methodology section.
17. Figure 1 – please show the full WRF domain and river catchments (in panelled plots) so that Figures 2, 6, 8, 10 are more easily understood. You cannot discern ice thickness differences in the figure, so consider removing.
18. Figure 2 – Without a clearer geographical setting, it is difficult to understand what is being shown here. What has been masked? A discrete color bar is needed, as well as a panel of topography.
19. Figure A1 – This is not useful. Include the catchment basins in Figure 1 and remove this figure.